GATE CSE

Resource Portal of GATE Overflow

Identify the Class of a Given Language

September 26, 2015 by Arjun Suresh (https://gatecse.in/author/arjunsuresh/)

Contents

Given a description of a language L, how to find the class of L?

Find the set of strings generated by L

1. If the set of strings in L is finite, L is regular since all finite languages are regular
2. If the set of strings in L is infinite, check if we can draw an NFA for recognizing L. If so, L is regular
3. If NFA is not possible for L, check if we can recognize L using a PDA, that is with a stack in addition to set of states. If so, L is CFL.
4. If the moves of PDA are all deterministic, then L is a $DCFL$
5. If PDA is not possible for L, see if we can get a TM for L
6. If TM takes only a linear space (in terms of length of input string), then L is CSL. otherwise its just recursive
7. If L is a decision problem and TM can just say “yes” and may not halt in case of “no”, then L is recursively enumerable (partially decidable).
8. If TM can say both “yes” and “no” then L is recursive
9. If no TM is possible for L, then L is undecidable

Some Facts

- Partially undecidable or semi-undecidable is considered undecidable. For example halting problem is considered undecidable but is semi-decidable.
- P, NP and NPC problems can all be decided by a TM and hence are

- Turing decidable problems are recursive but Turing recognizable (Turing acceptable) problems are only recursively enumerable.

Some Twisted Examples

1. $L = \{ww \mid w \in (a+b)^*\}$

The set of strings in L are $\{aa, bb, aaaa, abab, baba, bbbb, aaaaaa, \ldots\}$. We cannot accept these strings using an NFA. Now, even a PDA is not possible as once we store w on stack, it can only be read back in reverse order. Thus, we require 2 stacks to recognize L. Now, L can be accepted by a TM in linear space and hence L is CSL.

2. $L = \{ww \mid w \in (a+b)^+\}$

Same explanation as above, L is CSL.

3. $L = \{ww_R \mid w \in (a+b)^*\}$

ww_R can be accepted by a PDA and hence is CFL. But we need a $NPDA$ for this as there is no deterministic way to identify where w ends and w_R starts. $wcw_R, w \in (a+b)^*$ is accepted by a $DPDA$ and hence is $DCFL$.

4. $L = \{ww_R \mid w \in (a+b)^+\}$

Same explanation as above. L is CFL.

5. $L = \{wxw \mid w, x \in (a+b)^*\}$

L is regular since $L = \Sigma^*$, by making $x = (a+b)^*$ and $w = \epsilon$. i.e.; the set of

CSL. Here, we cannot do any reduction and hence there is no way to accept a string without checking w before c and w after c are the same which requires an LBA.

6. $L = \{wxw \mid w, x \in (a+b)^+\}$

L doesn't contain all strings in Σ^* as the strings like $abab$ are not contained in L. All words starting and ending in a or starting and ending in b are in L. But L also contains words starting with a and ending in b like $abbab, aabbbabaab$ etc where the starting sub-string exactly matches the ending sub-string and at least a letter separates them. To accept such strings we need a TM with linear space (this is at least as hard as accepting $ww, w \in (a+b)^*$), making L, a CSL.

7. $L = \{wxw_R \mid w, x \in (a+b)^*\}$

L is regular. Since, w can be ϵ and $x \in (a+b)^*$, making $L = \Sigma^*$. i.e.; the set of strings generated by L is $\{\epsilon, a, b, aa, ab, ba, bb, aaa, \ldots\} = \Sigma^*$

This language is different from the $L = \{wcw_R \mid w \in \{a,b\}^*\}$ which is clearly a DCFL. Here, we cannot do any reduction and hence there is no way to accept a string without checking that the string after c is the reverse of the string before c, which requires a DPDA.

8. $L = \{wxw_R \mid w, x \in (a+b)^+\}$

The set of strings in L are $\{aaa, aba, aaaa, aaba, abaa, abba, baab, \ldots\}$ i.e.; L contains all strings starting and ending with a or starting and ending with b and containing at least 3 letters. Moreover, L doesn't contain any other strings. Thus

9. $L = \{wxwy \mid w, x, y \in (a+b)^+\}$

For any string to be in L, the beginning part of the string (w) must repeat at some other point (between the second and last characters) of the string (next w). Since y is there at the end which can generate any string, we can make w as small as possible as per the given condition. So, w can be either a or b. We can thus write regular expression for L as $a(a+b)^+a(a+b)^+ + b(a+b)^+b(a+b)^+$

10. $L = \{xwyw \mid w, x, y \in (a+b)^+\}$

Similar explanation for example 9, except that instead of first character being a or b we have the last character. So, regular expression for L will be $(a+b)^+a(a+b)^+a + (a+b)^+b(a+b)^+b$

11. $L = \{wxyw \mid w, x, y \in (a+b)^+\}$

Here, w is coming at the beginning and also at the end. Unlike as in example 8 or 9, we cannot restrict w to be a or b as a string starting with a can end in b and still be in L- example $abaaab$, where $w = ab$ and $x, y = a$. In short we need to compare the substring at the beginning of the string with that at the end, making this a CSL.

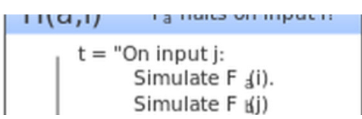
12. $L = \{xww \mid w, x \in (a+b)^*\}$

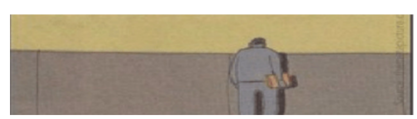
L is regular. Since, w can be ϵ and $x \in (a+b)^*$, making $L = \Sigma^*$. i.e.; the set of strings generated by L is $\{\epsilon, a, b, aa, ab, ba, bb, aaa, \dots\} = \Sigma^*$

14. $L = \{xww_R \mid w, x \in (a+b)^+\}$

Here, w cannot be ϵ and hence to accept the string we do need the power of a PDA making L a NCFL (non-determinism is required to guess the start of w).

Related





theorem/) /data_types_operators_in_c/) books-for-gatecse/)

Rice's Theorem with Examples (https://gatecse.in/rices-theorem/)
September 23, 2015
In "Theory of Computation"

C Book Chapter 2: Data Types and Operators in C (https://gatecse.in /data_types_operators_in_c/)
April 19, 2016
In "Programming & DS"

Best Books for GATE CSE (https://gatecse.in/best-books-for-gatecse/)
October 23, 2015
In "Books, Videos & Mock Tests"

Join the discussion…

LOG IN WITH

OR SIGN UP WITH DISQUS ?

Name

Chiranjeevi Kanaka • a year ago

epsilon is not present in any CSL, but in the 1st twisted example, it is given that the language is a CSL.

• Reply • Share ›

arjunsuresh1987 Mod → Chiranjeevi Kanaka • 3 months ago

epsilon is present in CSL although as an extension to the normal definition. Otherwise we cannot say CSL is a superset of CFL.

22 • Reply • Share ›

Sudeep Sharma • 2 years ago

Hello. The text is not formatting on mobile devices. Hope its not only my device.

• Reply • Share ›

Nipun • 2 years ago

In question 11, Language, L = {wxyw | x,y,w belonging to (a+b)^+ } should be regular as if w = abbab then , string in L = (abbab)(b)(a)(abbab) i.e. starting with a and ending with b. Similarly, if w =ba then, string will be of the form b(a+b)^+a.

Expression is all the strings starting with a and ending with b or all the strings starting with b and ending with a . Hence L is regular and NOT CSL.

• Reply • Share ›

Arjun Suresh Mod → Nipun • 2 years ago

ababbaba is in L or not?

• Reply • Share ›

GATE CSE Updates (https://www.facebook.com/gateoverflow/)

Google+ (https://plus.google.com/+GateoverflowIn)

Tweet (https://twitter.com/csegate)

FB Group (https://www.facebook.com/groups/gateoverflow)

Contact Us (https://gatecse.in/contact-us/)